NED

MULTINOMIAL LOGIT:
A LIMITED DEPENDENT VARIABLE TECHNIQUE

Michael LeBlanc
National Economics Divisions
Economics, Statistics, and Cooperatives Servic
U.S. Department of Agriculture
Washington, D.C. 20250

STAFF REPORT

NATIONAL
ECONOMICS
DIVISION

ECONOMICS,
STATISTICS
AND
COOPERATIVES
SERVICE

UNITED
STATES
DEPARTMENT OF
AGRICULTURE

This paper was prepared for limited distribution to the research community outside the U.S. Department of Agriculture. The views expressed herein are not necessarily those of ESCS or USDA.

MULTINOMIAL LOGIT: A LIMITED DEPENDENT VARIABLE TECHNIQUE

Michael LeBlanc
National Economics Divisions
Economics, Statistics, and Cooperatives Service
U.S. Department of Agriculture
Washington, D.C. 20250

September 1980

MULTINOMIAL LOGIT: A LIMITED DEPENDENT VARIABLE TECHNIQUE.

By Michael LeBlanc, National Economics Division; Economics, Statistics, and Cooperatives Service: U.S. Department of Agriculture, Washington, D.C. 20250. June 1980.

ABSTRACT

Many econometric analyses include dependent variables which are constrained to the interval between zero and one. Under such circumstances simple regression procedures break down. The simple logit model is extended to the multi-factor case. Two alternative models are defined depending on the error structure. The generalized least squares approach assumes the share specification is an accurate representation of the underlying input demand structure. The multinomial maximum likelihood model treats the dependent variable as a probability with a multinomial density. Either model provides a wide array of functional forms while maintaining a straight-forward estimating procedure.

KEYWORDS: Econometrics, Limited Dependent Variable, Multinomial Logit, Maximum Likelihood

MULTINOMIAL LOGIT: A LIMITED DEPENDENT VARIABLE TECHNIQUE

By

MICHAEL LEBLANC

SEPTEMBER 1980

I. INTRODUCTION

Many econometric analyses include dependent variables which are constrained to the interval between zero and one. Under these conditions simple regression procedures breakdown. Predictions may fall outside the zero one interval and, because of the grouped nature of the data, error terms are heteroscedastic.

The objective of this paper is to outline a multinomial logit methodology as a sensible response to the limited dependent variable problem and to provide a more robust alternative than translog cost function approaches to estimating derived demand.

The first section of the paper describes the limited dependent variable problem and the two approaches, probit and logit, typically used to resolve this difficulty. The simple logit model is extended to the multi-factor case and discussed in a probability model context. Two alternative models are suggested and estimating methodology for each is discussed. Finally, the implications for normalizing the models to facilitate estimation is addressed.

II. LIMITED DEPENDENT VARIABLE MODELS

Models with limited dependent variables have been used in many fields including biology, psychology, and economics. The earliest work was Berkson's (1944, 1951, 1953, 1955) applications to bioassay problems. The initial analysis in economics was Tobins (1958) automobile demand study. More recent applications using multinomial logit methodology are Tyrrell's

(1979) analysis of household expenditures, Baughman and Joskow's (1975) study of appliance choice, and Joskow and Mishkin's (1977) electric utility fuel choice model. In the case of a two variable probability choice model the simplest expression is

(1) $y_t = \alpha + \beta x_t + \varepsilon t$

where Y_t is a proxy for unobserved the probability P_t, x_t is an explanatory variable, and ε_t is the error term. The analyst's interest is the conditional probability of decision y_t, purchasing capital equipment, for example, given explanatory value x_t, the equipment price. It is not possible to know whether a particular firm will purchase equipment when the equipment's price is x_t. The objective of the analysis is to determine the probability of purchase. This implies a dependent variable confined to the interval from zero to one.

Simple linear regression equation (1) may result in prediction problems (See Figure 1).

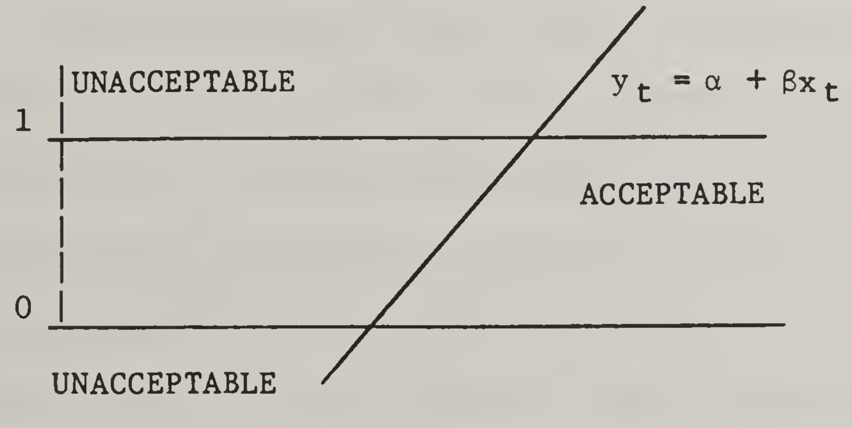

Figure 1. SIMPLE REGRESSION

While y_t is constrained between zero and one $\alpha + \beta x_t$ is unconstrained. It is quite possible, especially in a forecast mode, for predicted y to fall outside the zero-one bounds. In addition, since the residual

variance depends on the unknown probability value P_t the variances are heteroscedastic.

To mitigate the prediction problems a monotonic transformation is applied to expand the probability range from negative to positive infinity (See Figure 2).

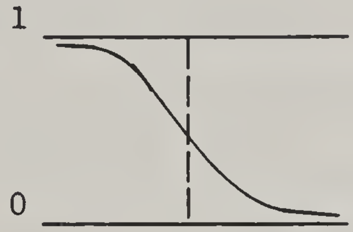

FIGURE 2. MONOTONIC TRANFORMATION OF A PROBATILITY TO THE RANGE $(-\infty, \infty)$.

The probability P is measured on the vertical axis and its transformation along the horizontal. As P increases from zero to one, its transformation increases from negative to positive infinity. The prediction problem of using a limited dependent variable is therefore avoided. There are many different transformations that may be applied. The most common transformations utilize either the normal distribution or logistic function.

<u>Probit</u>

By making a specific assumption about the distribution of the error term the parameters of equation (1) can be estimated. If the culmulative distribution function of a standardized normal variate is used the technique is termed probit analysis. The critical level of x triggering a purchase decision varies from observation to observation. By assuming the distribution of these critical levels is normal, the conditional probability, P(x), can be expressed. The dependent variable y is implicitly defined by

$$(2) \quad F(y) = P(x) = (1/2\pi)^{1/2} \int_{-\infty}^{\alpha + \beta x} \exp\left(\left(-\varepsilon/2\right)/2\right) d\varepsilon$$

where P is the probability. The transformation used is

$$y = F^{-1}(P) = \alpha + \beta x. \tag{3}$$

While circumventing the prediction difficulty the transformation has not mitigated the heteroscedasticity problem.

The transformed model's parameters can be estimated by maximum likelihood methods. Let $x_1, \ldots, x_n$ be observations for the independent variable. By arranging the observations where the first n* firms are purchasers and the last n - n* are not, the logarithmic likelihood function is

$$\ln L = \sum_{t=1}^{n^*} \ln P(x_t) + \sum_{t=n^*+1}^{n} \ln(1 - P(x_t)) \tag{4}$$

where

$$(1/2\pi\sigma)^{1/2} \exp(-(y_t - \alpha - \beta x_t)^2 / 2\sigma^2) \tag{5}$$

The maximum likelihood estimates are obtained by maximizing lnL with respect to α, β, and σ^2. The solution may be obtained using non-linear techniques.

Logit

Another transformation utilizes the logistic function. In a two state decision framework the logistic is written

$$Y_t = 1/(1 - \exp(g(x) + \varepsilon_t)) \tag{6}$$

where Y_t is defined above and $g(x_t)$ is an undefined function. By writing the odds in favor of equipment purchase as the ratio,

$$P/(1-P) = \exp(-(g(x) + \varepsilon)) \tag{7}$$

where P is the purchase probability, the range of the dependent variable is expanded from negative to positive infinity. The estimated form is obtained by taking the logarithm of equation (7)

$$(8) \quad \ln [P/(1-P)] = g(x) + \varepsilon.$$

The logarithm of the probability ratio is a monotonically increasing function of the probability P varing between negative and positive infinity. It is numerically equal to the logit of the complementary event but opposite in sign

$$(9) \quad \ln (P/(1-P)) = -\ln ((1-P)/P).$$

To estimate the logit function the error term is assumed to be generated by the binomial distribution. The odds or chance of success, equipment purchase, for any observation is analogous to sampling with replacement. The binomial distribution is

$$(10) \quad f(w) = \binom{n}{w} P^{w} (1-P)^{n-w}$$

where n is the total number of dollars, P is the probability of success, and w is the number of samples taken, share of equipment purchasers in the sample. If by specifying P_t as a logistic the likelihood function is

$$(11) \quad L = \prod_{t=1}^{n} \binom{n}{w} (1/(1 + \exp(-g(x_t))))^{w} \cdot (1 -(\exp g(x_t)/(1 + \exp(-g(x_t)))))^{n-w}$$

The density function f(w) is interpreted as the probability of choosing w positive responses from a sample of n.

The ability of the logistic function to constrain probability estimates to the range between zero and one is an important reason for the popularity of this model. In addition, the logit model is easy to estimate, especially when the logarithmic transformation is linear in the parameters. Furthermore, the underlying probability function gives a close approximation to the cumulative normal probability function.

III. MULTINOMIAL LOGIT

While the two state logit decision model provides an appropriate beginning, its usefulness is limited. Henri Theil (1969) first proposed a multinomial extension of the linear logit model to analyze consumer budget allocations. The multinomial logit is a flexible approach and for many applications provides a useful marriage between theoritically founded models and more ad hoc prediction oriented models. Unlike other share approaches for production or consumer behavior the multinomial logit does not require expenditures to be optimal or the result of a decision process having a unique mathematical representation. Flexability is manifested in the ability to incorporate many factors influencing expenditures. Since a major objective of this paper is to articulate an alternative to translog cost functions the following discussion is couched in terms of derived input demand.

The multinomial logit model indicates producer expenditures on factor inputs. A set of N independent conditional probabilities are assumed to jointly determine the allocation of expenditures into N categories. In addition, the probabilities are assumed to have the logistic structure and

are conditional on input prices and other independent variables. The probabilities are written

$$(12) \qquad \pi_i = \exp(g_i(P,x)) \Big/ \sum_{k=1}^{n} \exp(g_k(P,x))$$

where π_i is the conditional probability of a dollar being allocated to input i, P's are input prices, x's are other independent variables, and $g_i(P,x)$'s are generalized functions. No specific restrictions on $g_i(P,x)$ are necessary to apply the multinomial logit. However, to aid exposition the functions $g_i(P,x)$ are assumed linear in their parameters and invariant in structure between equations except for inter-equation parameter variability. The functions $g_i(P,x)$ are specified as

$$(13) \qquad g_i = \beta_{i1}h_1(P_1) + \beta_{i2}h_2(P_2) + \ldots + \beta_{iN}h_N(P_N) + \gamma_{i1}g_1(x_1) + \gamma_{i2}g_2(X_2) + \ldots + \gamma_{im}g_m(x_m) \qquad (i = 1,2,\ldots, N).$$

Since π_i is unobservable the model is made operational by using input expenditures or cost shares as a proxy. Equation (12) becomes

$$(14) \qquad S_i = P_iQ_i/M = \exp(g_i(P,x)) \Big/ \sum_{k=1}^{N} \exp(g_i(P,x)) \quad (i = 1,2,\ldots, N)$$

where

$$(15) \qquad M = \sum_{i=1}^{N} P_iQ_i.$$

To derive the sensitivity of the ith input share to changes in each independent variable the slope of the ith budget share with respect to any variable can be determined as

$$(16) \qquad S_i/\partial y = S_i(\partial g_i/\partial y - \sum_{j=1}^{N} S_j\ \partial g_j/\partial y) \qquad (i = 1,2,\ldots, N).$$

Caution is warranted when using estimated parameters from the multinomial logit model to derive demand price sensitivity. The logit approach is used to estimate parameters characterizing the underlying probability function. Therefore, using the parameters to characterize price sensitivity presumes input demand allocations are distributed in the same manner as the corresponding probabilities.

Elasticity

Two alternative derivations for price elasticities are possible. First, total expenditures are assumed invariant to input price changes. This is the production analogue to the budget allocation problem of consumer theory. The logit model is converted to a system of input demand equations

$$(17) \qquad Q_i = MS_i/P_i = \exp\ (f_i + \ln M - \ln P_i)/ \sum_{j=1}^{N} \exp(f_j) \qquad (i = 1,2,\ldots, N).$$

Given the "budget" assumption, the own-price elasticities, E_{ii}, cross-price elasticities, E_{ij}, and elasticities with respect to all non-price independent variables are derived for the ith good

$$(18) \qquad E_{ii} = P_i(\partial f_i/\partial P_i - \sum_{j=1}^{N} S_j\ \partial f_j/\partial P_i) - 1 \qquad (i = 1,2,\ldots, N)$$

$$(19) \qquad E_{ij} = P_j(\partial f_i/\partial P_j - \sum_{j=1}^{N} S_j \partial f_i/\partial P_j) \qquad (i = j = 1,2,\ldots, N)$$

$$(20) \qquad E_{ix_k} = x_k(\partial f_i/\partial x_k - \sum_{j=1}^{N} S_j \partial f_j/\partial x_k) \qquad (k = 1,2,\ldots, M).$$

The own-price elasticity approaches negative one and the cross-price elasticity approaches zero as the ith input share approaches one or the derivatives $\partial f_i / \partial P_i$ are invariant across shares. While the elasticities approach zero or one there are no a priori restrictions on either the magnitude or the sign of the elasticities.

If we assume, however, the total expenditure, M, is not invariant to changes in input prices and therefore $\partial M/\partial P_j \neq o$ then the own-price and cross-price elasticities are

$$(21) \quad E_{ij}{*} = w_i + P_i(\partial f_i/\partial P_i - \sum_{j=1}^{N} S_j \partial f_j/\partial P_i) - 1 \qquad (i = 1,2,\ldots, N)$$

$$(22) \quad E_{ij}{*} = w_i + P_j(\partial f_i/\partial P_j - \sum_{j=1}^{N} S_j \partial f_j/\partial P_j) \qquad (i = 1,2,\ldots, N).$$

In either case the calculation of E_{ix_k} remains the same. If the own-price elasticities are negative, the effect of input price increases on costs is to reduce price sensitivity by reducing total output.

IV. ESTIMATING THE MULTINOMIAL LOGIT

Specifying the multinomial logit input expenditure model requires selecting a functional form for f_i. The function f_i includes price and output effects and all other production or producer characteristics considered relevant. Assuming f_i is linear in parameters facilitates linear estimation techniques.

Two stochastic specifications of the model are possible. First, the share equation

$$(23) \quad S_i = P_i Q_i/M = \exp(f_i)/ \sum_{k=1}^{N} \exp(f_k) \qquad (i = 1,2,\ldots, N)$$

can be directly interpreted as a accurate characterization of producer expenditures for inputs. The stochastic model structure is generated by appending error terms ε_i to each function f_i. The input share equation, S_i, is transformed to an estimable form by normalizing with share N and taking logarithms. When f_i is linear in parameters estimated equations are

$$
\begin{aligned}
(24) \quad \ln(S_i/S_N) &= (\beta_{oi} - \beta_{on}) + (\beta_{ij} - \beta_{in})h(P_i) \\
&+ (\beta_{Ni} - \beta_{NN})h(P_N) + (\gamma_{oi} - \gamma_{oN}) \\
&+ (\gamma_{1i} - \gamma_{1N})g(x_1) + \ldots + (\gamma_{Ni} - \gamma_{NM})g(x_M) \\
&+ (\varepsilon_i - \varepsilon_N) \qquad\qquad (i = 1,2,\ldots, N).
\end{aligned}
$$

Clearly, the presence of ε_N in each equation generates non-zero covariance between equation error terms. Estimating each equation separately by ordinary least squares leads to inefficient parameter estimates. If each error, ε_i, is assumed homoscedastic, a generalized least squares estimator can be applied to improve estimator efficiency relative to an ordinary least squares procedure.

Zellner's (1962) genralized least squares technique for "Seemingly Unrelated Regressions" can be used to improve the efficiency of the estimators. If the same set of regressors is utilized in all N-1 equations then Zellner's procedure and ordinary least squares gives identical coefficient estimates. Estimates of the standard errors, however, differ.

Parameters estimated by the Zellner generalized least square procedure are not invariant to the choice of common denominator budget share when an estimated variance-covariance matrix is employed. Barten (1969), has shown maximum liklihood estimates are invariant to which equation is deleted and Kementa and Gilbert (1968) and Dhrymes (1970) show

iterating the Zellner procedure leads to maximum likelihood estimates. Therefore, iterating the Zellner procedure is a computationally efficient means of deriving maximum likelihood estimates. In general, the properties of maximum likelihood estimators only hold asymptotically. However, Kmenta and Gilbert (1968) suggest that most of the maximum likelihood estimator's asymptotic properties tend to be present in small samples.

The second interpretation of the multinomial logit is to assume S_i is a proxy variable for the unobservable probability π_i generated by the multinomial distribution

$$(25) \qquad f(x_1, x_2, \ldots, x_N) \;/\; (M!/(x_1!x_2! \ldots x_N!))\; P_1^{x_1} P_2^{x_2} \ldots P_N^{x_N}$$

where

$$(26) \qquad \sum_{i=1}^{N} P_i = 1.$$

The multinomial distribution characterizes the results of sampling with replacement from a population classified into N > 2 categories.* The sampling procedure can be viewed in several ways corresponding to different units of measure. The implication of different measurement unit is to alter the weights in the likelihood function. However, the maximum of each likelihood function is satisfied by the same set of probability estimates.

The multinomial expression utilized for the input demand expenditure system allocates each percentage of total expenditures independently. If shares are multiplied by 100 to facilitate the use of factorial notation then the multinomial distribution is

*See Fellner, W., "An Introduction to Probability Theory and Its Application," (John Wiley and Sons, 1968) p. 167.

$$(27)\qquad f(x) = (100!/(100S_1!100S_2! \ldots 100S_N!))\, \pi_1^{100S_1} \ldots \pi_N^{100S_N}$$

for period t.

Replacing π_i with the appropriate logit expression gives

$$(28)\qquad f(x) = (100!/(100S_1!100S_2! \ldots 100S_N!)).$$

$$(\exp(f_1)/\sum_{i=1}^{N}\exp(f_1))^{100S_1} \ldots (\exp(f_N)/\sum_{i=1}^{N}\exp(f_N))^{100S_N}$$

The joint multinomial distribution is used to generate a likelihood function of the form

$$(29)\qquad L = A^T \sum_{t=1}^{T} \sum_{i=1}^{N} (\exp(f_i)/\sum_{k=1}^{N} \exp(f_k))^{100S_{it}}$$

where

$$(30)\qquad A^T = \sum_{t=1}^{T} (100!/100S_{it}! \ldots 100S_{Nt}!).$$

Since the maximum likelihood estimator is invariant to monotonic transformations, the likelihood function is simplified by writing it in its logarithmic form

$$(31)\qquad \ln L = \text{Constant} + 100 \sum_{t=1}^{T} (\sum_{i=1}^{N} S_{it}f_{it} - \ln(\sum_{i=1}^{N} \exp(f_{it})))$$

The log-likelihood function is maximized where the first order derivatives with respect to the parameters and the matrix of second order derivatives is negative definite. The first order conditons are

$$(32)\qquad \partial \ln L/\partial \beta_{ik} = \sum_{t=1}^{T} x_{it}(S_{it} - \Pi_{it}) = 0 \qquad \begin{matrix}(i = 1,2,\ldots, N)\\(k = 1,2,\ldots, K).\end{matrix}$$

The second derivatives are therefore

$$\text{(33)} \qquad \partial^2 \ln L/\partial\beta_{ik}\, \partial\beta_{ik}' = \sum_{t=1}^{T} x_{kt}x_{kt}'\Pi_{it}(\Pi_{it} - 1)$$

where both parameters, β_{ik} and β_{ik} appear in the same equation and

$$\text{(34)} \qquad \partial^2 \ln L/\, \partial\beta_{ik}\, \partial\beta_{jk}' = \sum_{t=1}^{T} x_{kt}x_{kt}'\Pi_{it}\Pi_{jt}$$

and each parameter is from a different equation. Since the first-order conditions are non-linear an iterative search procedure, Newton-Raphson or Davidon-Fletcher-Powell for example, must be employed to solve for the maximum likelihood parameter estimates.

There is an indeterminancy in the equation system arising from the constraint

$$\text{(35)} \qquad \sum_{i=1}^{N} S_i = 1.$$

The share equations are invariant to the addition of the same expression, lnz for example, to every f_i since

$$\text{(36)} \qquad \exp(g_i + \beta \ln z)/ \sum_{j=1}^{N} \exp(f_j + \beta \ln z) =$$

$$\exp(g_i)\exp(\beta \ln z)/(\sum_{j=1}^{N} \exp(g_j)\ (\exp(\beta \ln z))$$

$$= \exp(g_i)/ \sum_{j=1}^{N} \exp(g_j) = S_i \qquad (i = 1,2, \ldots N).$$

This indeterminancy causes the matrix of second order partial derivatives in the multinomial maximum likelihood procedure to equal zero. The singularity problem is avoided by normalizing the N parameters for a particular variable in the N functions (Tyrrell, 1979). The normali-

zation imposed does not affect the computed elasticities or the predicted shares. The most straight forward approach is to divide the N-1 equations by the Nth equation. The equations are

(37) $S_i/S_N = \exp(f_i - f_N)$ $(i = 1,2, \ldots, N-1)$.

In the estimated logarthmic form the equations are

(38) $\ln (S_i/S_N) = (\beta_{oi} - \beta_{oN}) + (\beta_{1i} - {}_{iN})h_1(P_1) +$
$(\beta_{Ni} - \beta_{NN})h_N(P_N) + (\gamma_{oi} - \gamma_{oN}) + (\gamma_{1i} - \gamma_{iN})$
$g_1(x_1) + \ldots + (\gamma_{Ni} - \gamma_{NM})g_M(x_M)$ $(i = 1,2, \ldots, N)$.

The indeterminancy caused by the adding-up constraint prevents the calculation of the individual β_{ik}'s but allows the derivation of both the predicted probabilities and the elasticities. To facilitate discussion assume only price variables are included in the estimated model. Therefore, $h_i(P_i)$ equals P_i where P_i is the ith price. If the share system is normalized using the Nth share, then the estimated shares are

(39) $\ln S_i^* = \hat{\beta}_{io} + \hat{\beta}_{i1}P_1 + \ldots + \hat{\beta}_{iN} P_N$ $(i=1,2, \ldots, N-1)$

where S_i^* is equal to the predicted ratio of the ith to the Nth share and $\hat{\beta}_{ij}$ is equal to the estimated coefficient for the ith share and jth variable. The predicted Nth allocation is derived by

(40) $$\hat{S}_N = 1/(1 + \sum_{j=1}^{N-1} \exp(\ln S_i^*)\hat{S}_N))$$

since

(41) $$1 = 1/(\hat{S}_N + (\sum_{j=1}^{N-1} \exp(\ln S_i^*)\hat{S}_N).$$

The other N-1 allocations are easily calculated by

$$(42) \qquad S_i = S_N \exp(\ln S_i^*).$$

To derive price elasticities the Nth elasticity must necessarily be derived first. The predicted elasticity for the Nth allocation with respect to the ith price is

$$(43) \qquad \hat{\eta}_{Ni} = - P_i \sum_{j=1}^{N-1} \hat{S}_j \, \partial \hat{f}_j^*/\partial P_i + \begin{cases} -1 & \text{if } i=N \\ 0 & \text{otherwise} \end{cases}$$

$$(i=1,2,\ldots, N).$$

The ability to estimate the Nth rests on the equivalence of

$$(44) \qquad - P_i\left(\sum_{j=1}^{N-1} S_i \partial f_j^*/\partial P_i\right) = P_i\left(\partial f_j/\partial P_i - \sum_{i=1}^{N} S_j \partial f_j/\partial P_i\right)$$

where the right-hand side of equation (44) is derived from the unnormalized price elasticity formulation given in equations (18) (19), and (20). The two components in equation (44) are equivalent since

$$(45) \qquad - P_i\left(\sum_{j=1}^{N-1} S_j \, \partial f_j^*/\partial P_i\right) = - P_i\left(\sum_{j=1}^{N-1} S_j \, (\beta_{ji} - \beta_{Ni})\right)$$

$$= - P_i\left(\sum_{j=1}^{N-1} S_j \, \beta_{ji} - \beta_{Ni} \sum_{j=1}^{N-1} S_j\right)$$

$$= - P_i\left(\sum_{j=1}^{N-1} S_{ji} - \beta_{Ni}(1-S_N)\right)$$

$$= - P_i\left(\sum_{j=1}^{N-1} S_j \, \beta_{ji} - \beta_{Ni} + \beta_{Ni} \, S_N\right)$$

$$= P_i\left(\beta_{Ni} - \sum_{j=1}^{N} \beta_{ji} \, S_j\right).$$

$$(46) \qquad \hat{\eta}_{ji} = \hat{\eta}_{Ni} + P_i \partial f_j^*/\partial P_i \qquad (i=1,2, \ldots, N-1).$$

Therefore, although values cannot be estimated for each β_{ji} coefficient predicted shares and elasticities can be derived. The only restriction is that in each case, shares and elasticities, the Nth value is calculated first.

V. CONCLUSION

Limited dependent variable estimating techniques are receiving increased attention for consumer budget and derived demand input studies. The multinomial logit model provides a flexible alternative to the more popular translog expenditure approach by allowing a wide range of explanatory variables while constraining predictions to the interval between zero and one. Two general models are defined depending on the error structure. The generalized least squares approach assumes the share specification is an accurate representation of the underlying input demand structure. The multinomial maximum likelihood model treats the dependent variable as a probability with a multinomial density. Either model provides a wide array of functional forms while maintaining a straight-forward estimating procedure.

REFERENCES

Barten, A.P., "Maximum Likelihood Estimation of a Complete System of Demand Equations," European Economic Review Vol. 1 (Fall, 1969) pp. 7-73

Berkson, J., "Application of the Logistic Function to Bioassay, "Journal of the American Statistical Association Vol. 39 (September 1944) pp. 327-339

Berkson, J., "A Statistically Precise and Relatively Simple Method of Estimating the Bio-Assay with Quantal Response, Based on the Logistic Function," Journal of the American Statistical Association Vol. 48 (1953) pp. 565-599.

Christensen, L.R., D. Jorgenson, and L. Lau, "Transcendental Logarithmic Production Fontiers," The Review of Economics and Statistics Vol. 55 (February 1973) pp. 28-45.

Christensen, L.R., D. Jorgenson, and L. Lau, "Transcendental Logarithmic Utility Functions," American Economic Review Vol. 65 (June 1975) pp. 369-382

Dhrymes, P. "Equivalence of Iterative Aitken and Maximum Likelihood Estimates for a System of Regression Equations," Unpublished University of Pennsylvania, 1970

Fletcher, R. and M.J.D. Powell, "A Convergent Descent Method for Minimization," Computer Journal (1963) pp. 163-168

Kmenta, Jan and R. Gilbert, "Small Sample Properities of Alternativé Estimators of Seemingly Unrelated Regressions," Journal of the American Statistical Association Vol. 63 (December 1968) pp 1180-1200

Maddala, G.S., Econometrics (New York: McGraw-Hill Book Co., 1977)

McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," Chapter IV in Frontiers in Econometrics Vol. 1., P. Zarembka, ED. (New York: Academic Press, 1974) pp. 105-142

Oum, T.H., "A Warning on the Use of Linear Logit Models in Transport Mode Choice Studies," The Bell Journal of Economics Vol. 10 (Spring 1979) pp. 374-390

Theil, H., "A Multinomial Extension of the Linear Logit Model," International Economic Review Vol. 10 (October 1969) pp. 251-259

Theil, H. Principles of Econometrics (New York: John Wiley and Sons, Inc. 1971).

Tobin, J. "Estimation of Relationships for Limited Dependent Variables," Econometrica Vol. 26 (January 1958).

Tyrrell, T., "An Application of the Multinomial Logit Model to Predicting the Pattern of Food and Other Household Expenditures in the Northeastern United States," (Unpublished Ph.D. Dissertation, Cornell University, 1979)

Zellner, A. 'An Efficient Method for Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias," Journal of the American Statistical Association Vol. 57 (June 1962) pp. 348-368.

a